

Adaptively Optimizing the Algorithms for Adaptive Antenna Arrays for Randomly Time-Varying Mobile Communications Systems

Robert T. Buche and Harold J. Kushner
Division of Applied Mathematics
Brown University
Providence RI 02912

Abstract

Adaptive antenna arrays are widely used and have great promise to reduce the effects of interference and to increase capacity in mobile communications systems. Consider a single cell system with an (receiving) antenna array at the base station. The usual algorithms for obtaining the antenna weights for the adaptive array depend on parameters that are held fixed no matter what the operating situation, and the performance can strongly depend on the values of these parameters. For example, at time k , we might seek the antenna weights that minimize the performance function $E \sum_{l=1}^k \alpha^{k-l} e_l^2$, where e_l is the error in reception at sample time l . Typically, $\alpha < 1$ to allow tracking as conditions change. The performance of the algorithm for adapting the weights in the antenna array depends heavily on the chosen value of the forgetting or discount factor α . Generally, the optimal value will change rapidly in time as the operating conditions change. In some cases (for example, where the Doppler frequency of the mobile being tracked oscillates), the optimal value of α will also oscillate. We are concerned with the adaptive optimization of such parameters by the addition of another adaptive loop. The antenna weights and the value of α must be adapted simultaneously. We give an algorithm for adapting α , which is based on an approximation to a natural “gradient descent” method. The algorithm is practical and can improve the operation considerably. This is justified via simulations under a variety of operating conditions. The algorithm tracks the optimal value of α very well, and always performs better than the algorithm that uses any fixed α , sometimes much better. The adaptation can be based on a pilot signal or it can be partially blind. The adaptive algorithm for the parameter can be analyzed via stochastic approximation (SA) theory, where the SA algorithm is that for adapting α .

Methods Keywords: Stochastic Processes, Optimization, Control Theory.

I. INTRODUCTION

The adaptive antenna problem: Formulation. Adaptive antenna arrays are widely used and have great promise to reduce the effects of interference and to increase capacity in wireless systems [3], [8], [10]. The usual algorithms for adaptive arrays depend on parameters that are held fixed no matter what the operating situation, and the performance can strongly depend on the values of these parameters. We are concerned with the adaptive optimization of such parameters by the addition of another adaptive loop. It will be seen that the method is practical and can improve the operation considerably.

We consider the problem of optimizing reception at the base station of a single cell system with r antennas. The updates of the antenna weights are to be done in discrete time, but there are natural continuous time analogs. Let $\bar{x}_{i,k}, i \leq r$, denote the complex (baseband) output of antenna i at measurement time k ; it is the sum of the signals due to all of the mobiles, plus additive noise. Let $\bar{w}_{i,k}$ denote the complex weight assigned to antenna i at time k . Define the vectors $\bar{X}_k = \{\bar{x}_{i,k}, i \leq r\}$, $\bar{w}_k = \{\bar{w}_{i,k}, i \leq r\}$. Let $\{s_k\}$ denote a real-valued pilot training sequence from the particular mobile that is being tracked. It is assumed that the training sequence is known at the receiver. The algorithm to be presented also works well with partially blind adaptation, or with only periodic use of a known

Partially supported by Army Research Office Contract DAAD19-00-1-0549 and National Science Foundation Grant ECS 9979250

Partially supported by Army Research Office Contract DAAD19-00-1-0549 and National Science Foundation Grant ECS ECS 0097447

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2002		2. REPORT TYPE		3. DATES COVERED 00-00-2002 to 00-00-2002	
4. TITLE AND SUBTITLE Adaptively Optimizing the Algorithms for Adaptive Antenna Arrays for Randomly Time-Varying Mobile Communications Systems			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Brown University, Division of Applied Mathematics, 182 George Street, Providence, RI, 02912			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 12	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

pilot signal, with blind adaptation in between. The weighted output of the array is $\Re\{\sum_i \bar{w}_i^* \bar{x}_{i,k}\} = \Re\{\bar{w}_k^* \bar{X}_k\}$, where \Re denotes the real part. Henceforth, to simplify the notation, we concatenate the real (\Re) and imaginary (\Im) components. Thus, for complex \bar{x}_k , let the unbarred quantity x_k denote the concatenation ($\Re \bar{x}_k, \Im \bar{x}_k$) and define X_k, w_k, etc , analogously.

The algorithm for adapting the weights. For $\alpha \in (0, 1]$ and fixed weight vector w , define the discounted cost

$$J_k(\alpha, w) = \sum_{l=1}^k \alpha^{k-l} |s_l - w' X_l|^2. \quad (1.1)$$

Typically $\alpha < 1$ to allow tracking of changing circumstances. Suppose that we wish to minimize the performance function $E J_k(\alpha, w)$ over w , for large k . Let w_k denote the value of w that minimizes $J_k(\alpha, w)$. Define the errors $e_k(w) = s_k - w' X_k$ and $e_{k+1} = s_{k+1} - w'_{k+1} X_{k+1}$. The standard recursive least squares algorithm for computing w_k can be written as [7]

$$\begin{aligned} w_{k+1} &= w_k + L_{k+1} e_{k+1}, \\ L_{k+1} &= \frac{P_k X_{k+1}}{\alpha + X'_{k+1} P_k X_{k+1}}, \\ P_{k+1} &= \frac{1}{\alpha} \left[P_k + \frac{P_k X_{k+1} X'_{k+1} P_k}{\alpha + X'_{k+1} P_k X_{k+1}} \right], \end{aligned} \quad (1.2)$$

where the initial weight w_0 and matrix P_0 are given. The discounted pathwise cost, given the sequence of optimal weights $\{w_l\}$, is

$$J_k(\alpha) = \sum_{l=1}^k \alpha^{k-l} e_l^2. \quad (1.3)$$

The results of simulations of this algorithm for mobile communications were reported in [10].

On the value of α . The value of the performance function (equivalently, the error rates) can be quite sensitive to the value of the discount factor α , as will be seen when the numerical data is presented at the end of the paper. We are concerned with the question of its optimal value. If all of the mobiles are stationary and the variance of the additive noise constant, then the optimal value of α will be unity. If the mobiles (particularly the one being tracked) are moving rapidly and the additive noise level is small, then the optimal value of α will be relatively small. In practice, the optimal value might vary rapidly, perhaps changing significantly many times per second, as the operating conditions change. Our aim is the development of a practical algorithm for adapting α ; i.e for tracking its current optimal value. It is based on an intuitively reasonable gradient descent idea. The simulations that are presented in Section V show the rapid response of the adaptive procedure, as well as the impressive gains in performance that can be achieved.

Outline of paper. In the next section we give some background material concerning an adaptive algorithm for the problem of tracking time-varying parameters in a linear system, via noisy observations. Although it will not be used in the sequel, the essential features of that problem are also those of the problem of interest here, since a balance must be found between the averaging of the noise effects (i.e., large α) and the ability to track (i.e., small α). The adaptive procedure finds the proper balance. There, as in our problem, there are two levels of adaptation. One is that which estimates the parameters (equiv. our antenna weights), and the other adapts a parameter in the first adaptive algorithm to optimize its performance. The success of that time-varying parameter tracking algorithm was the motivation for the one used here, even though the quantities being tracked are different and our current problem is more complicated. In Section III, we define the adaptive procedure which is to be used for our problem. Then, in Section IV, the precise model that defines the $x_{i,k}$ will be given. The simulation results in Section V clearly demonstrate

the utility of the approach and behavior of the algorithm under a variety of challenging operating conditions, with different INR (interference to noise) and SINR (signal to interference plus noise) ratios. The algorithm is a form of stochastic approximation with “state-dependent” noise [6]. Its behavior can be well approximated by the solution of a mean ODE (ordinary differential equation), and some brief comments on the mathematical arguments which provide the theoretical justification that the algorithm is of the gradient descent type are at the end of Section III.

II. BACKGROUND: A PARAMETER TRACKING ALGORITHM

In this section, we will discuss a simple form of an adaptive algorithm for a related problem, which motivated the actual algorithm to be used here. Let y_n be a member of a stationary real-valued random sequence and ϕ_n a (stationary) r -dimensional random vector that is correlated with it. The r -dimensional vector θ that minimizes the mean square error $E|y_n - \theta' \phi_n|^2$ is $\bar{\theta} = [E\phi_n \phi_n']^{-1} E\phi_n y_n$. A stochastic approximation algorithm for recursively estimating $\bar{\theta}$ is

$$\theta_{n+1}^\epsilon = \theta_n^\epsilon + \epsilon \phi_n [y_n - \phi_n' \theta_n^\epsilon], \quad (2.1)$$

where $\epsilon > 0$ is some chosen step size parameter.

Consider, for the moment, the special form where $y_n = \sum_{i=0}^{r-1} a_i(n) \chi_{n-i} + \rho_n$, where the $\{\chi_n, \rho_n\}$ are mutually independent and each of the χ_n and ρ_n are identically (in n) distributed and $E\rho_n = 0$. Define $\phi_n = (\chi_n, \dots, \chi_{n-r+1})$. If the $a_i(n)$ do not depend on n , then $\bar{\theta} = \{a_0, \dots, a_{r-1}\}$. Now suppose that y_n is still stationary, but the parameters $\bar{\theta}_n = \{\alpha_0(n), \dots, \alpha_{r-1}(n)\}$ do vary with time. Then we wish to track $\bar{\theta}_n$ via an algorithm of the form (2.1). If $\bar{\theta}_n$ varies rapidly, then the step size ϵ should be large, to allow tracking. If it varies slowly, but the variance of the χ_n or ρ_n is large, then ϵ should take a small value. In general, not only do we not know what the optimal value of ϵ is, but it changes over time. We need to adapt the value of ϵ , based on the measurements, and at the same time that $\bar{\theta}_n$ is being estimated.

Consider the following procedure for choosing the estimates ϵ_n of the best current value of ϵ . Fix $\epsilon_n = \epsilon$. Define the error $e_n(\epsilon) = y_n - \phi_n' \theta_n^\epsilon$. The scheme suggested in [1, p. 160] and developed in [2], [5] is to find the value of ϵ that minimizes the stationary value

$$E[y_n - \phi_n' \theta_n^\epsilon]^2 / 2 = Ee_n^2(\epsilon) / 2. \quad (2.2)$$

Formally, let V_n^ϵ denote the “derivative” $\partial \theta_n^\epsilon / \partial \epsilon$ for the stationary process. The random variable θ_n^ϵ is not a classical function of ϵ , although its distribution depends on ϵ . But, it can be shown [5] that the asymptotic values of the V_n^ϵ as used in what follows can be interpreted as mean square derivatives. Formally differentiating (2.2) with respect to ϵ , assuming stationarity, and setting the derivative to zero yields

$$0 = -E[y_n - \phi_n' \theta_n^\epsilon] \phi_n' V_n^\epsilon = -Ee_n(\epsilon) \phi_n' V_n^\epsilon. \quad (2.3)$$

Dropping the E on the right side of (2.3) yields the “stochastic estimate” $-e_n(\epsilon) \phi_n' V_n^\epsilon$ of the gradient of (2.2) with respect to ϵ at time n . Formally differentiating (2.1) with respect to ϵ yields

$$V_{n+1}^\epsilon = V_n^\epsilon - \epsilon \phi_n \phi_n' V_n^\epsilon + \phi_n [y_n - \phi_n' \theta_n^\epsilon].$$

This discussion suggests the following algorithm for adapting θ and ϵ . Let $\mu > 0$ be small and, for $e_n = y_n - \phi_n' \theta_n$, use the algorithm

$$\theta_{n+1} = \theta_n + \epsilon_n \phi_n e_n, \quad (2.4)$$

$$\begin{aligned} \epsilon_{n+1} &= \max \{0, \epsilon_n + \mu e_n \phi_n' V_n\}, \\ V_{n+1} &= V_n - \epsilon_n \phi_n \phi_n' V_n + \phi_n [y_n - \phi_n' \theta_n], \quad V_0 = 0. \end{aligned} \quad (2.5)$$

The algorithm (2.4) and (2.5) has two levels. The parameter $\bar{\theta}_n$ is tracked by (2.4), while the optimal step size parameter ϵ is tracked by (2.5). The actual stochastic approximation algorithm is the one for ϵ_n in (2.5), where the step size is μ . The sequence $\{\phi_n, y_n, V_n\}$ plays the role of a driving noise process. The dependence of V_n on $\{\epsilon_n\}$ and

$\{\theta_n\}$ is quite complicated. It is what we call “state-dependent noise” [6]. In applications, the performance is a great deal less sensitive to the value of μ in (2.4) than the original algorithm (2.1) is to the choice of ϵ . Despite the apparent formality in the derivation, the algorithm performs well with ϵ_n being nearly optimal under broad conditions. It is a significant improvement over (2.1), with its fixed value of ϵ . A proof of convergence and supporting simulations are [5].

III. THE ADAPTIVE ALGORITHM FOR α

For $w_k \equiv w$, the expression (1.1) for $J_k(\alpha, w)$ can be written in the recursive form

$$J_{k+1}(\alpha, w) = \alpha J_k(\alpha, w) + e_k^2(w).$$

If the values of w_k are determined by the least squares algorithm (1.2), then we can write

$$J_{k+1}(\alpha) = \alpha J_k(\alpha) + e_k^2. \quad (3.1)$$

If the value of α changes with time, then we are concerned with the discounted performance function EJ_k , where

$$J_{k+1} = \alpha_k J_k + e_k^2. \quad (3.2)$$

Suppose, for the moment, that the value of α is fixed and that $\{X_k, w_k, s_k\}$ is stationary, where w_k is determined by (1.2). The stationary distributions will depend on α . For this process, define

$$J_k(\alpha) = \sum_{l=1}^k \alpha^{k-l} e_l^2.$$

Suppose that that we wish to choose α to minimize the stationary expectation $EJ_k(\alpha, w)$ over α for large k . The time-varying parameter tracking algorithm of Section II was based on a “derivative,” or “mean-square derivative.” In our case, the dependence of the algorithm (1.2) on α is quite complicated. To avoid dealing with the rather messy forms that would result from a differentiation of the right sides of (1.2) with respect to α , we simply work with a finite difference form. Typically, the function $\bar{J}(\alpha) = \min_w EJ_\infty(\alpha, w)$ is strictly convex and continuously differentiable. In our model, the value increases sharply as α increases beyond its optimal value, and increases more slowly as α decreases below its optimal value. It is somewhat insensitive to α around the optimal value. Finite difference estimators, for the difference intervals that we use, provide excellent approximations.

Let $\delta > 0$ be a small difference interval, let α_k denote the value of α at the k th update, and define $\alpha_k^\pm = \alpha_k \pm \delta/2$. The algorithm (1.2) is run for both α_k^\pm . Thus, we define the two sets of recursions, for $+$ and $-$:

$$\begin{aligned} e_{k+1}^\pm &= s_{k+1} - [w_k^\pm]' X_{k+1}, \\ w_{k+1}^\pm &= w_k^\pm + L_{k+1}^\pm e_{k+1}^\pm, \\ L_{k+1}^\pm &= \frac{P_k^\pm X_{k+1}}{\alpha_k^\pm + X_{k+1}' P_k^\pm X_{k+1}}, \\ P_{k+1}^\pm &= \frac{1}{\alpha_k^\pm} \left[P_k^\pm + \frac{P_k^\pm X_{k+1} X_{k+1}' P_k^\pm}{\alpha_k^\pm + X_{k+1}' P_k^\pm X_{k+1}} \right]. \end{aligned} \quad (3.3)$$

For small $\mu > 0$, the adaptive algorithm for α is

$$\alpha_{k+1} = \alpha_k - \mu \frac{[e_{k+1}^+]^2 - [e_{k+1}^-]^2}{\delta}. \quad (3.4)$$

The initial weight w_0 was found using the least-squares solution ($\alpha = 1$ in (1.1)) from an initial small block of data. The initial matrix P_0 is just the inverse of the sample covariance matrix using data from this block. Initial parameter $\alpha_0 \in (0, 1)$ is arbitrarily chosen.

If α is fixed, and the (temporarily assumed) stationary distribution is used, then for small $\delta > 0$, the (stationary) expectation of the coefficient of μ in (3.4) should be close to the derivative of $EJ_k(\alpha)$ with respect to α . The intuitive idea behind the algorithm is that the value of α changes much more slowly than that of w , so that we are essentially in the stationary state. In this case, we clearly have a stochastic algorithm driven by a process whose values are estimates of the negative of a gradient. It turns out that the idea can be justified, both in simulations (Section V) and mathematically. There are more equations to work with in the finite difference form, as opposed to the form using formal derivatives. But the equations are much simpler. Various simpler stochastic approximation forms were also dealt with, and will be discussed elsewhere.

Comment on the asymptotic properties of the algorithm. Using results from the theory of stochastic approximation, one can analyze the algorithm (3.3) and (3.4). The basic stochastic approximation algorithm is (3.4), since μ is small. The quantities $(X_k, w_k^\pm, s_k, L_k^\pm, P_k^\pm)$ play the role of noise. In this paper, we are concerned with the simulations and the presentation and motivation of the algorithm. There is little space for dealing with the convergence, and we confine ourselves to a few motivational remarks. Owing to the way that the evolution of this noise is tied to that of α_k we have what is called state-dependent noise [6, Chapter 8]. The asymptotic behavior (small μ) of (3.4) is determined by a mean ODE, whose right hand side is a “local” average of the coefficient of μ in (3.4). Loosely speaking, since the α_k sequence varies much more slowly than do the driving noises, one can compute this local average by assuming that α_k is fixed. Suppose that there is a $\bar{g}(\alpha)$ and m such that

$$\frac{1}{m} \sum_{l=n}^{n+m-1} E_n \frac{[e_l^-]^2 - [e_l^+]^2}{\delta} \approx \bar{g}(\alpha)$$

for large n, m , and α_l is held fixed at α , and where E_n is the expectation given the data to time n . Then the mean ODE is $\dot{\alpha} = \bar{g}(\cdot)$ [6, Chapter 8].

In our case, the process X_k will rarely be stationary or ergodic. For example, see (4.3) which is summed over j , i.e. the mobiles, to give \bar{X}_k . In (4.3), the dominant effect is that of the Doppler frequencies $\omega_{j,k}^d$. This frequency will change over time. But, over short time intervals, X_k varies rapidly in a periodic way and the desired averaging will occur. Full details for such problems are in [6, Chapter 8], and will be explored in detail for the current problem elsewhere. But, the overall conclusion is that, for small μ , the algorithm for adapting α behaves as a (finite difference approximation to a) gradient descent algorithm, which is what we were aiming for.

IV. THE PHYSICAL MODEL

In order to keep the simulations simple and focus on the essential issue of adaptation, the mobiles move in two dimensions. The three antennas are evenly spaced with spacing $d > \lambda/2$, where λ is the carrier wavelength (the carrier frequency is 800×10^6 Hz). The sample (bit) period is $h = 4 \times 10^{-5}$ sec. The number of strong interfering mobiles is either one ($N_I = 1$) or three ($N_I = 3$). Their amplitudes (i.e. square root of power) vary from about 1/4 that of the desired mobile, to being roughly equal. In the presented simulations, there is no scattering associated with these strongly interfering mobiles: i.e. they are in a line-of sight (LOS) environment. However, we note that when scattering is added, the effects of the adaptation are as impressive as the LOS case and the behavior of the adapted α is much more complicated. We model the effect of additional interferers with uniform scattering (Rayleigh fading) by adding complex-Gaussian noise to each antenna which is independent in time and across the antennas. This noise can be assumed to be independent across the antenna elements since $d > \lambda/2$. We assume a median field strength model where the signal amplitude at the receiver from mobile j at time k is $1/d_{j,k}^2$, where $d_{j,k}$ is the distance to a reference antenna in the array [4].

We assume a narrowband signal (carrier frequency \gg signal bandwidth) so the signal does not change appreciably over the time that it takes to traverse the antenna array. The interferers are in the far field so their transmitted electromagnetic wave can be assumed to be a plane wave at the antenna array. The pilot signal s_k , for the tracked or desired user, is assumed known. It is i.i.d., binary (+1, -1), and is independent of the signals from the other mobiles.

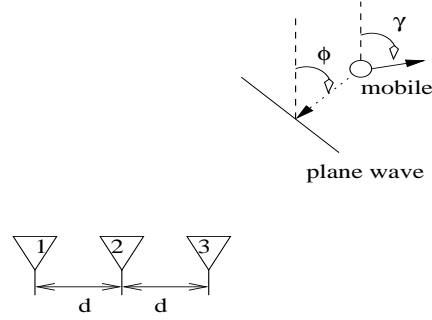


Fig. 1.

In practice, there would be either a training period or reference signals sent periodically, as part of the desired users synchronization signal.

The signal in each antenna is the sum of those emanating from the mobiles, plus complex Gaussian noise. The noise terms are assumed to be mutually independent, the real and complex parts are independent, and each has the same variance σ^2 . Define, as usual,

$$\text{SINR} = 10 \log \frac{P_{des}}{\sum_{i=1}^{N_I} P_i + 2\sigma^2}$$

and

$$\text{INR} = 10 \log \frac{\sum_{i=1}^{N_I} P_i}{2\sigma^2},$$

where P_{des} and P_i , resp., are the signal powers (at the antenna) of the desired and i th interfering mobile, resp. The most important factor in the determination of the optimal value of α at any time is the Doppler shift, although the values are also affected by the SINR and INR.

The Doppler frequency of mobile j at sample time k is

$$\omega_{j,k}^d = -\frac{2\pi}{\lambda} v_{j,k} \cos(\phi_{j,k} - \gamma_{j,k}), \quad (4.2)$$

where $\gamma_{j,k}$ is the angle of the travel of mobile j (see Figure 1), $v_{j,k}$ its speed, and $\phi_{j,k}$ the angle of arrival of its plane wave, all at sample time k . The spatial signature corresponding to a plane wave arriving at angle ϕ to the normal to the plane of the antennas (see Figure 1) is given by the column vector (antenna 1 is the reference antenna)

$$c = \left[1, \exp\left(-i \frac{2\pi}{\lambda} d \sin \phi\right), \exp\left(-i \frac{2\pi}{\lambda} 2d \sin \phi\right) \right]', \quad (4.1)$$

where λ is the carrier wavelength. We denote the spatial signature corresponding to mobile j at time k by $c_{j,k}$ where $\phi_{j,k}$ is used in (4.1). The component of the received signal at the antenna array at sample time k and which is due to mobile j is given by

$$\bar{x}_{j,k} = \frac{1}{d_{j,k}^2} s_{j,k} \exp\left(i \left[\sum_{l=1}^k \omega_{j,l}^d h \right]\right) c_{j,k}, \quad (4.3)$$

where h is the time interval, between updates (which is 4×10^{-5} seconds in our simulations). The (complex) signal received by the array at sample time k is $\bar{X}_k = \sum_j \bar{x}_{j,k}$. Of particular interest is the case where the wave number

$2\pi/\lambda$ is very large so that small variations in the mobility of the mobile can lead to large changes in the Doppler frequency. The signal X_k can be based either on TDMA or CDMA. In the latter case, it is measured after the matched filters (which use the signature of the desired user).

Note that the model is close to what was used in [10], except that we allow somewhat larger Doppler frequencies.

V. DISCUSSION OF THE SIMULATIONS

We will describe the performance of the algorithm for adapting α , via a set of simulations. Unless otherwise noted, the direction and velocity of each mobile evolved as a semi-Markov process, each moving independently of the others. They were constant for a short random interval, then there was sudden acceleration or deceleration in each coordinate, and so forth. In the plots, only the associated piecewise constant Doppler shifts are given, since that is the most important factor in the adaptation of α . The simulations start with two “baseline” sets of data, one for $N_I = 1$ and one for $N_I = 3$. Then the data is varied to explore the behavior under a variety of operating situations. For $N_I = 1$, the baseline is SINR = 5.3db, INR = 1.3 db, and for $N_I = 3$ it is SINR = 0.5db, INR = 7.9db. Mobile 2 is always the desired one. In each baseline case, the signal amplitude at the receiver of the interfering mobiles are approximately the same, and were approximately one fourth that of the desired mobile. This represents a large interference, especially with CDMA. The given SINR and INR values are only approximations, since the signal strength changes over the simulation interval; however, these changes are relatively small. We used $\mu = .0008$ and $\delta = .002$. Changing the value of μ up or down by factor of four had little effect on the overall performance.

Starting with each baseline case (described below), the data was varied systematically, as follows.

1. **Interferer(s) are moved closer to the antennas.** This results in an increased INR and a decreased SINR. The interferers are such that their power at the antennas is approximately the same as that of the desired mobile.
2. **Interferer(s) are moved further away from the antennas.** This results in a decreased INR and an increased SINR.
3. **Increased variance of the additive complex noise.** This corresponds to increasing the number of scattering sources. It results in a decreased SINR and a decreased INR.
4. **Decreased variance of the additive complex noise.** This results in an increased SINR and an increased INR.
5. **Other mobility models.** To get a better idea of the behavior of the adaptive algorithm, models where the Doppler frequency moved either in a “straight-line” or zigzagged were simulated.

We will see that α_k “tracks” the Doppler frequency of the desired user in all cases, and that there is a significant improvement in the performance, over that corresponding to the use of constant values of α . The performance is much less sensitive to the value of μ than it is to the value of α , which supports the conclusions in [5].

A. Baseline Cases

In all cases, mobile 2 is the desired one. First we note that if the mobiles are stationary, where $\alpha = 1$ is optimal, then the values of α_k in the adaptive procedure remained close to unity, and the sample mean square errors were virtually indistinguishable with those for the case $\alpha_k \equiv 1$.

The behavior of the adapted α -process as well as of the Doppler frequencies of a typical simulation for the baseline case with $N_I = 1$ are in Figure 2. Note that the Doppler shift is the vertical scale times 10^4 . Mobile 2 starts with a high Doppler frequency (corresponding to a velocity of approximately 150 km/hr), which then decreases suddenly at $t = .6$ sec., then decreases more slowly, and finally increases slightly. The behavior of α is typical for all simulations with this mobility for mobile 2. It initially oscillates about $\alpha = .85$, which is very close to the optimal value for the associated Doppler shift. Then, when the Doppler frequency drops to about 1.5×10^4 , α increases quickly, and then continues to increase (on the average) as the Doppler frequency continues to drop. At $t = 1$, the Doppler frequency rises slightly and then remains constant. Except for the brief transient periods, the values of α are close to the optimal. When smaller μ is used, the paths of α are smoother, the transient period longer, but the overall performance is very similar. Note that the behavior of the Doppler frequency of the interfering mobile had negligible affect on α .

Figure 3 plots the running mean squared error (or moving average [MA]) cost for the adapted algorithm together with those for the algorithm with constant values of α . At time k the MA cost is given by

$$\frac{1}{k} \sum_{l=1}^k e_l^2.$$

In our simulations we do not model the detection since we are focusing simply on the benefit of the adaptive- α algorithm. But, we note that in general $e_l^2 > 0$ when there is perfect detection so that our MA costs may seem high. The important point is the relative performance between the constant- α and adaptive- α algorithms. In Figure 3, the constant value $\alpha = .84$ gives results that are very close to the optimum. But, with other system data the best constant values are different, and the cost might still be significantly larger than for that of the adaptive algorithm. The use of constant values of α never outperformed the adaptive algorithm. Except for the cases of very high Doppler frequencies, the performance was approximately the same if blind adaptation were used, with the pilot signal being used only for initialization.

Figure 4 gives the adaptation process for the baseline case $N_I = 3$. The results are similar, despite the fact that the number of mobiles is greater than the number of antenna elements. Again, the behavior of the interfering mobiles had little effect on the evolution of α .

When the additive noise is increased, the optimal value of α increases. See Figure 5, where SINR=-1.5 and INR = -8.7db, which represents a large increase in the noise. The wilder behavior of α is due to the larger noise. There is a fairly short term memory in these algorithms, so the randomness in the noise sequence has a significant affect on α . The behavior is smoother if smaller μ or larger δ is used. But, in all cases, the adaptive algorithm outperformed the constant α forms, sometimes significantly (see Figure 6). If the variance of the additive noise is decreased, then the optimal values of α decrease, and the adaptive algorithm still kept close to the optimal value. When the noise variance is smaller, the other properties of the paths of the desired and interfering mobiles play a greater role, although the dominant influence is still the Doppler frequency of the desired mobile.

Moving the interfering mobiles further out or closer in had little effect on the paths of the adapted α , although the paths jumped about the mean values somewhat more for the smaller INR cases. For $N_I = 3$, and the mobiles further out, SINR=7.1, INR=-2.4db. A typical plot of the α -path is in Figure 7.

For an example where the Doppler frequency of the desired mobile zigzagged in a saw-tooth fashion, see Figure 8, for the case $N_I = 3$. The optimal value of α also varies in a saw-tooth fashion, and the adaptive algorithm tracks them very well. For another example, where the Doppler frequency of the desired mobile is linearly decreasing, see Figure 9 for the case $N_I = 1$. The adaptive algorithm tracks very well. These examples also illustrate the property that the cost is better (sometimes much better) than the alternative of using any fixed value of α (see Figure 10 for the saw-toothed example).

REFERENCES

- [1] A. Benveniste, M. Metivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximation*. Springer-Verlag, Berlin and New York, 1990.
- [2] J.-M. Brossier. *Egalization Adaptive et Estimation de Phase: Application aux Communications Sous-Marines*. PhD thesis, Institut National Polytechnique de Grenoble, 1992.
- [3] G.J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, autumn 1966.
- [4] W.C. Jakes. *Microwave Mobile Communications*. Wiley-Interscience, New York, 1974.
- [5] H.J. Kushner and J. Yang. Analysis of adaptive step size SA algorithms for parameter tracking. *IEEE Trans. Automatic Control*, AC-40:1403-1410, 1995.
- [6] H.J. Kushner and G. Yin. *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, Berlin and New York, 1997.
- [7] L. Ljung. *System Identification Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1986.
- [8] G.V. Tsoulos, Editor. *Adaptive Antenna Arrays for Wireless Communications*. IEEE Press, New York, 2001.
- [9] B. Widrow, P.E. Mantey, L.J. Griffiths, and B.B. Goode. Adaptive antenna systems. *Proc. IEEE*, 55:2143-2159, December 1967.
- [10] J.H. Winters. Signal acquisition and tracking with adaptive arrays in digital mobile radio system IS-54 with flat fading. *IEEE Trans. on Vehic. Techn.*, 42:377-393, 1993.

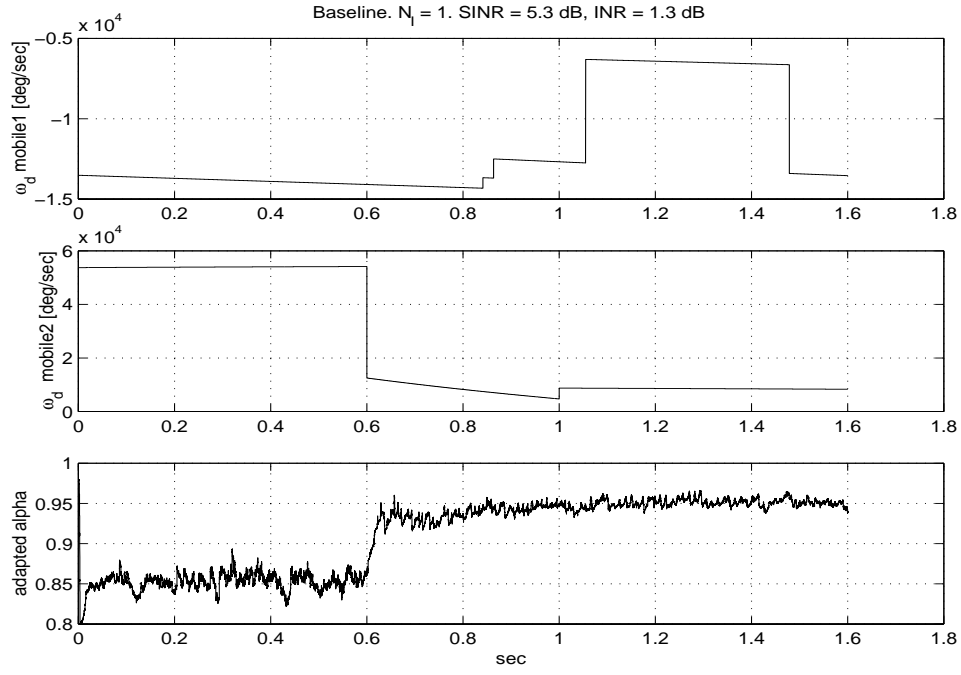


Fig. 2.

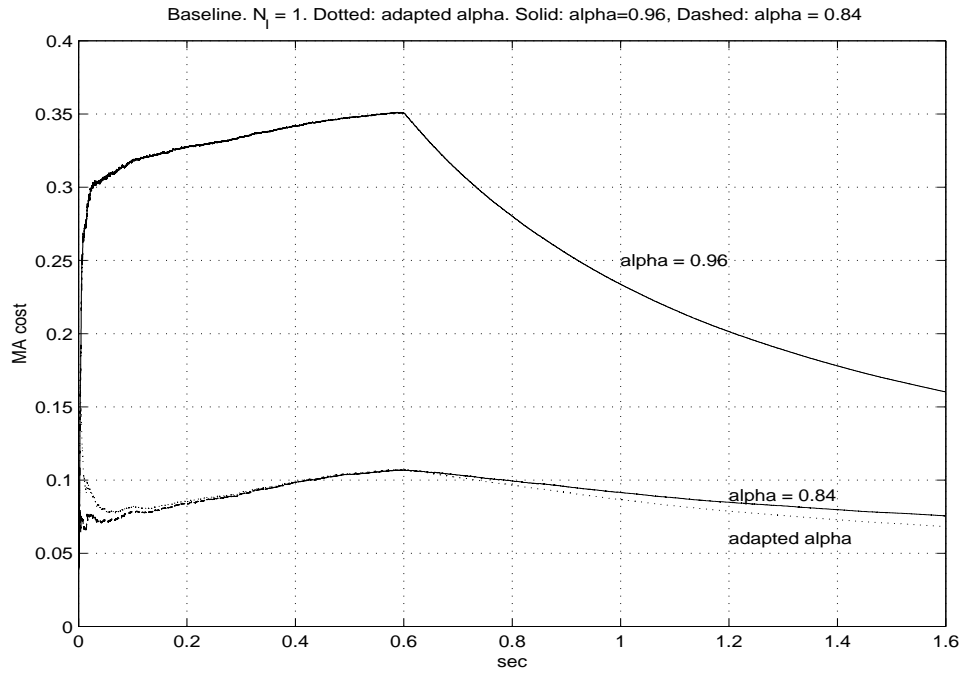


Fig. 3.

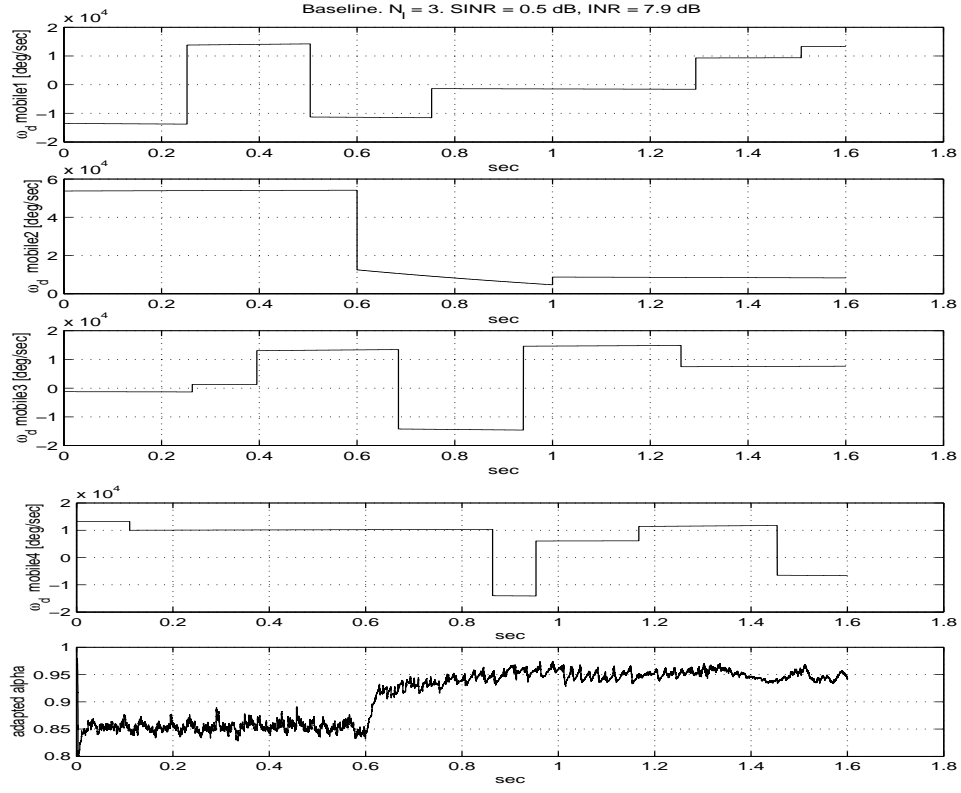


Fig. 4.

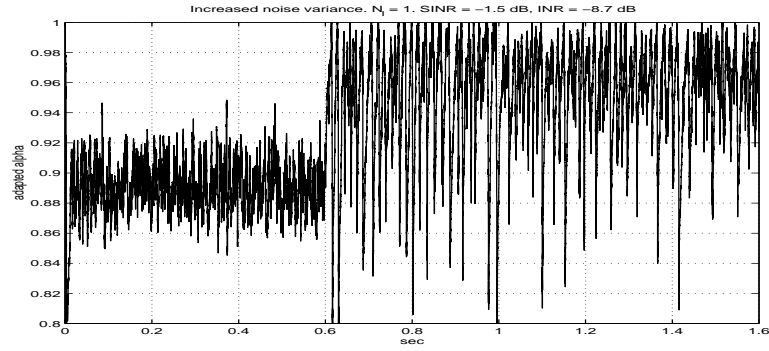


Fig. 5.

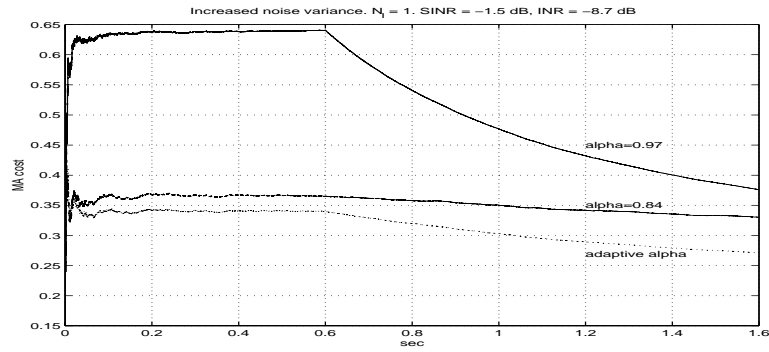


Fig. 6.

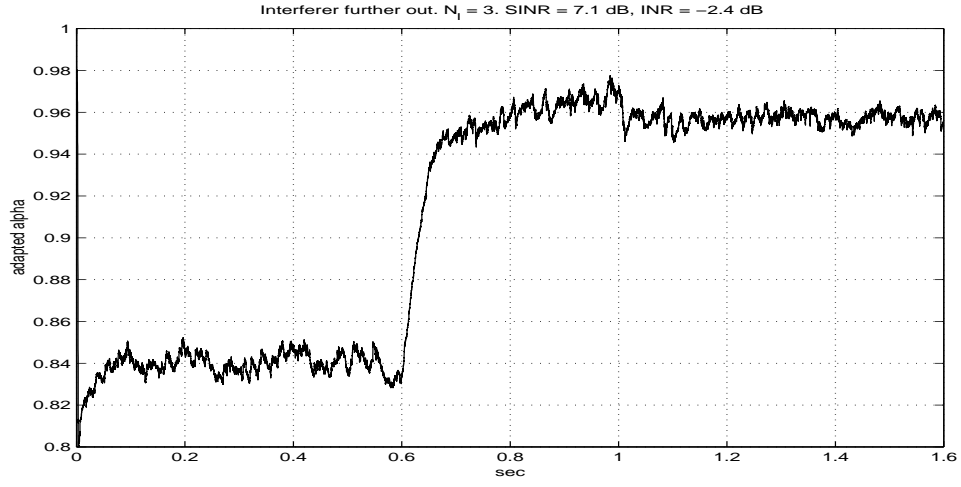


Fig. 7.

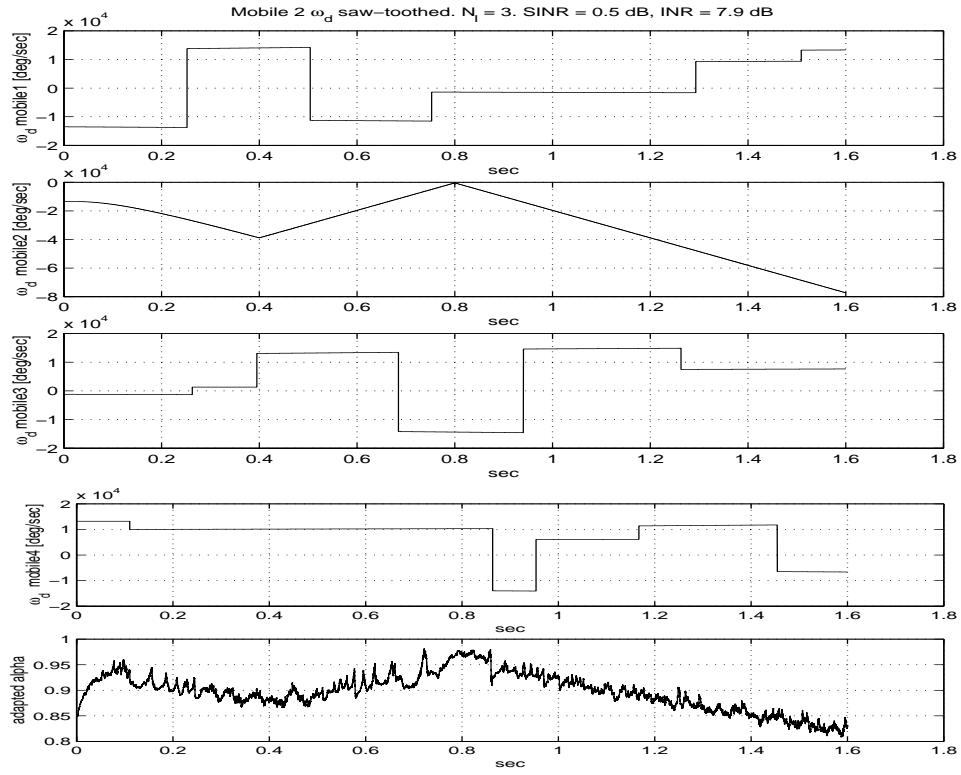


Fig. 8.

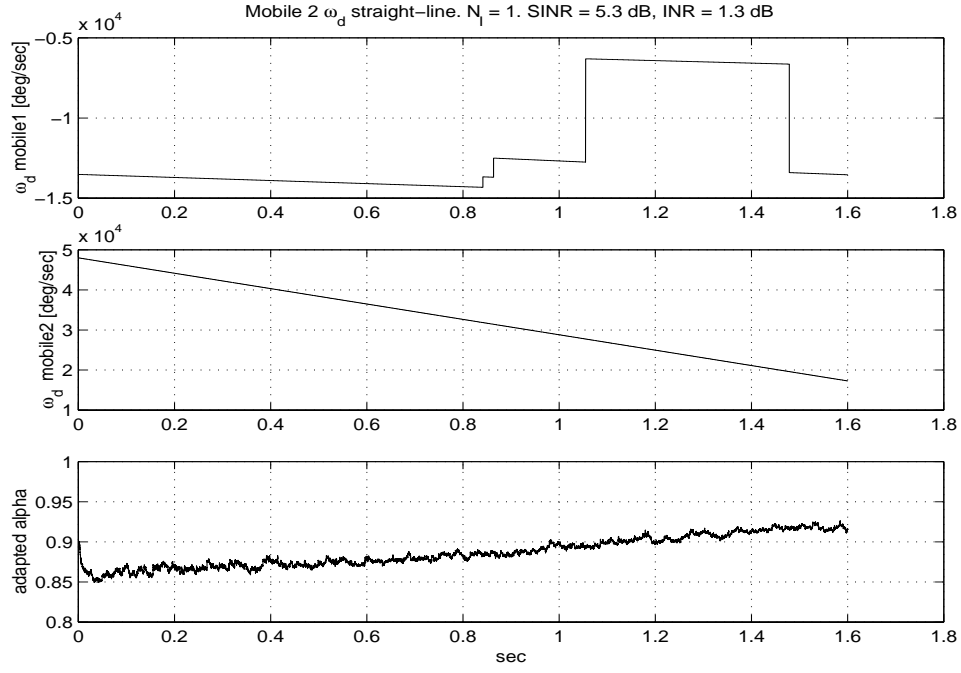


Fig. 9.

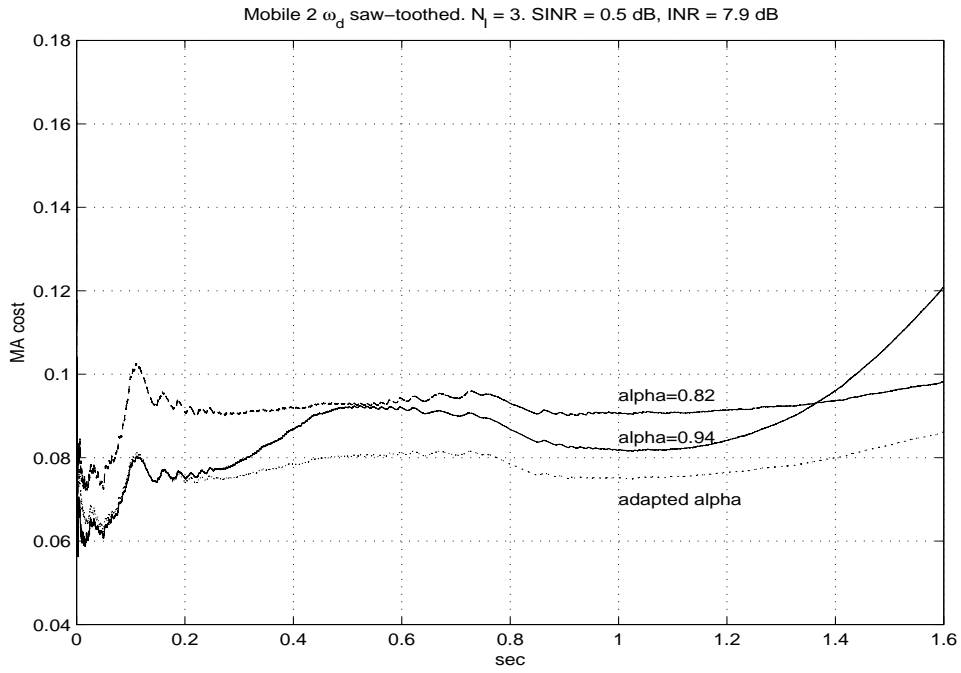


Fig. 10.